Journal of Blood Medicine

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Identification of Novel *F9* Gene Variants in 143 Vietnamese Patients with Hemophilia B

Khanh Quoc Bach[1–3], Chinh Quoc Duong[1,2], Huong Thi Bich Vu[1], Binh Thanh Ngoc Nguyen[1], Trang Thuy Nguyen[1], Mai Thi Nguyen[1], Ruoxin Li[4], Wendy Hutchison[4], Farisha Shabnam Esaq[4], Huyen Tran[5], Thanh Ha Nguyen[1,3]

[1]National Institute of Hematology and Blood Transfusion, Hanoi, Vietnam; [2]University of Medicine and Pharmacy, Vietnam National University, Hanoi, Vietnam; [3]Hanoi Medical University, Hanoi, Vietnam; [4]Thalassemia and Haemophilia Molecular Reference Laboratory, Monash Medical Centre, Melbourne, Victoria, Australia; [5]Haemophilia Treatment Centre, Australian Centre for Blood Diseases, Monash University, Melbourne, Victoria, Australia

Correspondence: Thanh Ha Nguyen, National Institute of Hematology and Blood Transfusion, No. 5 Pham van Bach Street, Cau Giay District, Hanoi, Vietnam, Email nguyenhathanh@hmu.edu.vn

Purpose: Vietnam is estimated to have approximately 30,000 hemophilia B (HB) carriers, with hundreds of new cases registered annually. However, comprehensive molecular studies on HB remain limited. Therefore, this study aimed to characterize genetic variants and assess their clinical significance in unrelated Vietnamese patients with HB.

Patients and Methods: This study included a cohort of 143 unrelated HB patients with diagnosed FIX levels. Genetic analysis of the *F9* gene was performed using DNA sequencing and other molecular techniques. Variant pathogenicity was classified following ACMG/AMP guidelines, supplemented by computational predictions and clinical data.

Results: A 100% variant detection rate was achieved, identifying 83 unique variants from 143 patients. Single nucleotide variants were predominant, with missense variants accounting for 71.08%. Of the 83 unique variants, 20 novel variants were identified, including six missenses, four nonsenses, four frameshifts, two large deletions, two in-frame deletions, and two splice-site variants. The serine protease domain contained the highest proportion of variants (49.4%). Pathogenicity analysis revealed a predominance of severe phenotypes (72.03%). Among the novel variants, twelve were classified as pathogenic, one as likely pathogenic, and seven as variants of uncertain significance. A noteworthy case was the NM_000133.4:c.-21C>T promoter variant associated with HB Leyden, which demonstrated age-dependent improvements in factor IX levels.

Conclusion: This study expands the mutational spectrum of HB in the Vietnamese population and provide critical insights into genotype-phenotype correlations. The identification of novel variants enhances diagnostic precision and underscores the importance of comprehensive genomic analyses in understanding disease mechanisms.

Keywords: factor IX, bleeding disorder, blood disease, next-generation sequencing, genetic disease, coagulation factor

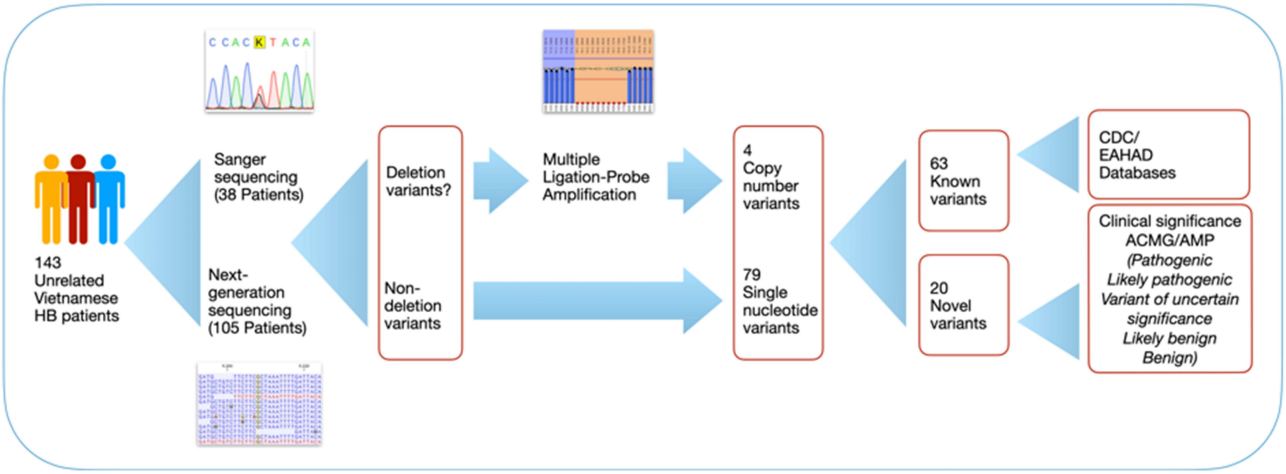
Introduction

Hemophilia B (HB) is a rare X-linked recessive bleeding disorder caused by variants in the *F9* gene, leading to a deficiency or dysfunction of coagulation factor IX (FIX). HB is further classified into mild, moderate, and severe forms based on the residual plasmatic activity of FIX. This condition predominantly affects males, while females typically act as carriers. Symptomatic cases in females are rarely reported.[1] The global birth prevalence of HB is approximately 1 in 30,000 male births, although reported rates vary considerably between countries, ranging from 0.01 to 8.07 per 100,000 males.[2] These variations likely reflect differences in diagnostic capabilities and healthcare systems.[3]

In Vietnam, the epidemiology of HB reflects global trends and local factors. The estimated prevalence of HB in the Vietnamese population is approximately 1 in 60,000 males. Data from the Hemophilia Center at the National Institute of Hematology and Blood Transfusion, Vietnam's central Hemophilia center, suggests that hundreds of new cases are

Received: 25 December 2024
Accepted: 18 March 2025
Published: 1 April 2025

Graphical Abstract



identified each year.[4] Furthermore, the center estimates that around 30,000 individuals in the country are carriers of the condition.[4]

The *F9* gene, located on the X chromosome (Xq27.1), spans approximately 34 kilobases (kb) and comprises eight exons and seven introns. The molecular characterization of *F9* gene variants has shown remarkable progression over the past three decades, with considerable expansion in the documentation of unique mutations. Initial report in 1990 by Giannelli et al identified 216 unique variants, establishing the foundation for systematic mutation cataloging in HB.[5] A pivotal expansion occurred with the establishment of an interactive database platform, which by 2013 had documented over 3,000 pathogenic mutations and polymorphisms.[6] As of 2024, the database included more than 3,700 variants (Factor IX Variant Database, http://www.factorix.org/), reflecting enhanced molecular diagnostic capabilities and increased genetic screening in previously underrepresented populations. This exponential growth in variant identification has significantly improved the understanding of genotype-phenotype correlations and facilitates more accurate genetic counseling and therapeutic decision-making in HB management. Understanding DNA variations in the *F9* gene is essential for establishing diagnostic protocols, facilitating prenatal diagnosis, and enhancing the care, management, and treatment of HB.

However, Vietnam currently lacks comprehensive studies on the *F9* gene and HB as a specific bleeding disorder. The identification of novel *F9* variants holds substantial clinical significance for advancing HB patient care in Vietnam's healthcare landscape. In the Vietnamese context, where healthcare resources require optimal allocation, understanding specific genetic variants helps prioritize high-risk patients and guides prophylaxis regimens. The capability to conduct prenatal screening and genetic counseling becomes more robust with expanded variant knowledge, enabling informed reproductive decisions for affected families. Furthermore, novel variant identification contributes to the global understanding of *F9* mutational spectrum, particularly significant for Asian populations, which may harbor distinct genetic profiles. This molecular insight supports Vietnam's progression toward precision medicine approaches in HB management, potentially reducing treatment complications and improving cost-effectiveness through targeted interventions. The integration of novel variant data into clinical practice ultimately strengthens the foundation for evidence-based HB care in Vietnam's developing healthcare system. Therefore, this study aimed to investigate the prevalence and clinical significance of genetic variations associated with HB in the Vietnamese population. By elucidating the genetic landscape of this disorder, this study seeks to enhance the understanding of HB and provide valuable insights to support its management and treatment in Vietnam.

Material and Methods

This study included a cohort of 143 unrelated patients diagnosed with HB following the criteria of the World Federation of Hemophilia. All patients were managed at the Hemophilia Center of the National Institute of Hematology and Blood Transfusion.

Peripheral blood samples collected from the HB patients were used for coagulation assays and subsequent genetic analyses. FIX levels were measured using a standard 1-stage clotting assay with HemosIL FIX-deficient plasma (Instrumentation Laboratory Company, MA, USA). Inhibitor screening was conducted systematically in all patients using the activated partial thromboplastin time (APTT) mixing test without incubation at 6-month intervals. Additional screening assessments were initiated upon detection of suboptimal therapeutic response to FIX replacement therapy. In cases where the APTT mixing study demonstrated incomplete correction, inhibitor quantification was subsequently performed utilizing the Bethesda assay methodology. For molecular analysis, genomic DNA was extracted from peripheral blood using the QIAmp DNA Blood kit (QIAGEN, Germany).

Genetic Analysis Techniques

Genetic variant analysis was adapted from protocols established by the Mayo Clinic and the UK Haemophilia Genetics Laboratory Network.[7,8] DNA analysis of the *F9* gene was performed using Next-generation sequencing for 105 samples and Sanger sequencing for 38 samples. For patients with a suspected large gene deletion/insertion, multiplex ligation-dependent probe amplification (MLPA) was performed to assess the sample.

Primer Design and F9 Gene Amplification

NGS primers were designed using the Primer3 tool,[9] targeting the complete DNA sequence of the *F9* gene (NM_000133.4), including all exons, introns, and untranslated regions. In silico evaluation of primers was performed using the Nucleotide BLAST tool (NCBI) to ensure precise targeting of the *F9* DNA sequence. Primer binding temperatures (Tm) were optimized using gradient PCR. Long-range amplification method was utilized to successfully amplify large fragments of *F9* gene, using UltraRun LongRange PCR master mix (QIAGEN, Germany) and optimized PCR program. This includes the first denaturation step at 95 °C for 3 minutes, then 35 cycles of 95°C for 30 seconds followed by 65°C for 5 minutes, the last step of 65°C for 10 minutes was added to complete the target amplification.

Sanger sequencing primers were followed the protocol from the Thalassemia and Haemophilia Molecular Reference Laboratory, Monash Medical Centre, Melbourne, Australia. The nucleic acid sequences of primers are listed in Supplement 1.

Molecular Analysis

Sanger sequencing was performed on the AB3500 Genetic Analyzer (ThermoFischer, USA) using BigDye Terminator v3.1 Mix (ThermoFishcher, USA) and DyeEx 96 kit (QIAGEN, Germany). For NGS analysis, DNA fragments in the target gene region were enriched and indexed using Nextera XT DNA Samples Preparation and Nextera XT Index Kits (Illumina, USA) and then sequenced on the MiSeq sequencer with the MiSeq v2 Reagent Kit – 300 cycles (Illumina, USA). DNA libraries were subjected to paired-end sequencing with 152 cycles from each end. A minimum coverage depth of 30X was required for accurate variant identification. Samples with missing PCR amplicons or undetected variants were analyzed for the DNA copy number variation using SALSA MLPA Probemix P207 F9 (MRC Holland, The Netherlands).

Sequencing data analysis and variant calling were performed using CLC Genomics Workbench version 20 (QIAGEN, Germany), while MLPA data was interpreted using Coffalyser software (MRC Holland, The Netherlands). The cDNA (c). and protein (p). nomenclature of *F9* variants were based on the RefSeq reference sequence database (NCBI) NM_000133.4 and protein reference sequence NP_000124.1 according to the recommendations of the Human Genome Variation Society (HGVS). Variant classification was based on the 2015 American College of Medical Genetics and Genomics guidelines for sequence variant interpretation.

Variant Interpretation

Variant databases from the European Association for Hemophilia and Allied Disorders and the US Centers for Disease Control and Prevention were referenced to confirm the presence of pathogenic variants in the patients. The interpretation of novel variants in the *F9* gene was conducted using the ACMG/AMP guidelines supplemented with the quantitative scoring framework outlined.[10,11] For in silico variant prediction analysis, distinct computational tools are employed based on variant type: REVEL score for missense variants, SPLICE AI for splice site alterations, and SIFT INDEL for insertions and deletions. Moreover, each variant was evaluated for its potential pathogenicity by applying weighted criteria across evidence categories, including population frequency, computational predictions, functional studies, and clinical presentation. Loss-of-function variants, such as nonsense, frameshift, and critical splice site changes, were assigned strong evidence under PVS1, while population rarity (PM2), evolutionary conservation (PM1), and computational predictions (PP3) provided additional support. Phenotypic data from patients were integrated as supporting evidence (PP4) when available. Using the point-based scaling of ACMG/AMP, the total evidence was quantified to assign variants into one of five classification tiers: pathogenic, likely pathogenic, uncertain significance, likely benign, or benign.

Results

Characteristics of the Study Group

The study cohort comprised 143 unrelated individuals, including one female patient. The female case resulted from her mother being a carrier and her father being an HB patient. The mean age of the cohort was 23 years (range: 3–76 years). Most patients (72.03%) exhibited severe disease phenotypes, while only a few (4.9%) presented with mild disease. No patients were detected to have developed FIX inhibitors. More details are shown in Table 1.

Identification of F9 Gene Variant in Patients with Hemophilia B

DNA analysis confirmed that all patients carried variants in the *F9* gene, yielding a 100% detection rate. Eighty-three unique variants were identified and distributed across coding, non-coding, and untranslated regions. Most of the variants (n = 75) were located in the coding regions spanning exons 1–8. Exon 8 had the highest number of variants (n = 35), including missense and frameshift mutations, followed by exon 2 with 13 variants. Two variants were observed in the 5′ UTR region, while six variants were detected in intronic regions. These intronic variants included splice-site alterations such as NM_000133.4:c.277+4A>G and NM_000133.4:c.278–3A>G. There was one case with a variant at position c.-21C>T in the promotor region of the *F9* gene (NM_000133.4). This patient initially had prepubertal FIX levels of 4%, which increased to 8.9% at age 14 and 23.6% at age 18.

Table 1 Characteristics of the Study Group

Characteristics		Number of Patients (%)
FIX level (%)	< 1% (severe)	103 (72.03)
	1–5% (moderate)	33 (23.08)
	> 5%–< 40% (mild)	7 (4.90)
FIX Inhibitor	With inhibitor	0
	Without inhibitor	143 (100)
Sex	Male	142 (99.30)
	Female	1 (0.70)
Age (years)	23±16 (min: 3; max:76)	

Abbreviation: FIX, Factor IX.

Most of the variations were single nucleotide variants (SNVs), predominantly missense mutations (71.08%, n = 59). Nonsense mutations and splice-site variations each accounted for 6.02% (n = 5) of the total variants. Notably, large deletions affecting one to eight exons constituted approximately 6% of all variants, distributed as follows: single-exon deletion (1.21%), two-exon deletion (2.41%), three-exon deletion (1.21%), and complete gene deletion (1.21%). Small deletions were observed in 4.42% of cases, equally distributed between single codon deletions (2.41%) and frameshift mutations (2.41%).

The mutations exhibited diverse patterns across different functional domains, with the serine protease domain containing the highest number of variants (49.39%, n = 41). The GLA domain contained nine variants (10.84%), while the EGF domains collectively had 12 variants (14.46%), distributed between EGF1 (n = 6) and EGF2 (n = 6).

Severity classification demonstrated a clear predominance of severe phenotypes, with 66 variants (79.52%) associated with severe disease manifestation. Nine variants accounted for moderate phenotypes (10.85%), while mild expressions were observed in five cases (6.02%). Notably, three variants (3.61%) exhibited mixed severity patterns.

Detailed characteristics and distribution of the variant are summarized in Table 2 and Figure 1.

Table 2 Summary of the 83 Unique F9 gene Variants Identified in 143 Unrelated Vietnamese Patients with Hemophilia B

No	HGVS cDNA	HGVS Protein	Exon/Intron	Domain	Number of Patients	Severity (Clinical Data)	Reported
1	c.-55G>A	N/A	5'UTR	Promoter	2	S	Yes
2	c.-21C>T	N/A	5'UTR	Promotor	1	Mo	No
3	c.88G>T	p.(Val30Phe)	Exon 1	PRO	1	S	No
4	c.128G>A	p.(Arg43Gln)	Exon 2	PRO	4	S, Mo.	Yes
5	c.127C>T	p.(Arg43Trp)	Exon 2	PRO	7	S, Mo, M.	Yes
6	c.137G>T	p.(Arg46Met)	Exon 2	PRO	1	Mo	Yes
7	c.137G>A	p.(Arg46Lys)	Exon 2	PRO	1	S	Yes
8	c.149G>A	p.(Gly50Asp)	Exon 2	GLA	3	M	Yes
9	c.161_162del	p.(Glu54Valfs*7)	Exon 2	GLA	1	S	Yes
10	c.190T>C	p.(Cys64Arg)	Exon 2	GLA	1	S	Yes
11	c.192T>G	p.(Cys64Trp)	Exon 2	GLA	1	S	No
12	c.206G>A	p.(Cys69Tyr)	Exon 2	GLA	3	S	Yes
13	c.212T>C	p.(Phe71Ser)	Exon 2	GLA	1	M	Yes
14	c.215A>G	p.(Glu72Gly)	Exon 2	GLA	1	S	Yes
15	c.223C>T	p.(Arg75*)	Exon 2	GLA	4	S	Yes
16	c.224G>A	p.(Arg75Gln)	Exon 2	GLA	1	M	Yes
17	c.253–1G>C	N/A	Intron 2	N/A	1	S	Yes
18	c.277+4A>G	N/A	Intron 3	N/A	1	S	Yes
19	c.278–3A>G	N/A	Intron 3	N/A	2	S	Yes
20	c.344A>G	p.(Tyr115Cys)	Exon 4	EGF1	1	S	Yes
21	c.348_349insGA	p.(Cys117Aspfs*15)	Exon 4	EGF1	1	S	No
22	c.349T>A	p.(Cys117Ser)	Exon 4	EGF1	2	S	Yes
23	c.355T>C	p.(Cys119Arg)	Exon 4	EGF1	1	S	Yes

(*Continued*)

https://doi.org/10.2147/JBM.S514338

Table 2 (Continued).

No	HGVS cDNA	HGVS Protein	Exon/Intron	Domain	Number of Patients	Severity (Clinical Data)	Reported
24	c.382T>C	p.(Cys128Arg)	Exon 4	EGF1	1	S	Yes
25	c.383G>A	p.(Cys128Tyr)	Exon 4	EGF1	2	S	Yes
26	c.418-?_864+?del	p.(?)	Exon 4 to 6	EFG1-EFG2-AP	1	S	No
27	c.407_408insGTAACAT	p.(Asn135*)	Exon 5	EGF2	1	S	No
28	c.422G>A	p.(Cys141Tyr)	Exon 5	EGF2	1	S	Yes
29	c.433T>C	p.(Cys145Arg)	Exon 5	EGF2	1	S	Yes
30	c.470G>T	p.(Cys157Phe)	Exon 5	EGF2	1	S	No
31	c.547-?_749+?del	p.(?)	Exon 5	EGF2	1	S	No
32	c.518_520del	p.(Ala173del)	Exon 5	EGF2	1	S	Yes
33	c.520+13A>G	N/A	Intron 5	N/A	1	Mo	Yes
34	c.520+2T>C	N/A	Intron 5	N/A	1	S	No
35	c.536G>A	p.(Gly179Glu)	Exon 6	Linker	1	S	Yes
36	c.571C>T	p.(Arg191Cys)	Exon 6	Linker	1	Mo	Yes
37	c.577G>T	p.(Glu193*)	Exon 6	Activation	1	S	No
38	c.655C>T	p.(Gln219*)	Exon 6	Activation	1	S	Yes
39	c.676C>T	p.(Arg226Trp)	Exon 6	Activation	2	S	Yes
40	c.677G>A	p.(Arg226Gln)	Exon 6	Activation	2	S	Yes
41	c.689G>T	p.(Gly230Val)	Exon 6	SP	1	S	No
42	c.718T>C	p.(Trp240Arg)	Exon 6	SP	1	S	Yes
43	c.755G>C	p.(Cys252Ser)	Exon 7	SP	1	S	Yes
44	c.761G>A	p.(Gly254Asp)	Exon 7	SP	1	S	Yes
45	c.865-?_2800+?del	p.(?)	Exon 7 to Poly A	SP	1	S	Yes
46	c.838+1G>A	N/A	Intron 7	N/A	1	S	Yes
47	c.871G>A	p.(Glu291Lys)	Exon 8	SP	1	Mo	Yes
48	c.874C>T	p.(Gln292*)	Exon 8	SP	1	S	No
49	c.880C>T	p.(Arg294*)	Exon 8	SP	3	Mo, S.	Yes
50	c.881G>A	p.(Arg294Gln)	Exon 8	SP	18	Mo	Yes
51	c.914A>G	p.(Tyr305Cys)	Exon 8	SP	1	S	Yes
52	c.929delA	p.(Asn210Ilefs15*)	Exon 8	SP	1	S	No
53	c.942T>A	p.(His314Gln)	Exon 8	SP	1	M	Yes
54	c.947T>C	p.(Ile316Thr)	Exon 8	SP	1	S	Yes
55	c.952del	p.(Leu318Phefs*8)	Exon 8	SP	2	S	No
56	c.956T>C	p.(Leu319Pro)	Exon 8	SP	3	S	Yes
57	c.1004G>A	p.(Cys335Tyr)	Exon 8	SP	1	S	Yes

(*Continued*)

Table 2 (Continued).

No	HGVS cDNA	HGVS Protein	Exon/Intron	Domain	Number of Patients	Severity (Clinical Data)	Reported
58	c.1024A>C	p.(Thr342Arg)	Exon 8	SP	1	S	Yes
59	c.1025C>T	p.(Thr342Met)	Exon 8	SP	3	Mo	Yes
60	c.1060del	p.(Ser354Valfs*14)	Exon 8	SP	1	S	No
61	c.1069G>C	p.(Gly357Arg)	Exon 8	SP	1	S	No
62	c.1084A>T	p.(Lys362*)	Exon 8	SP	1	S	No
63	c.1088G>A	p.(Gly363Glu)	Exon 8	SP	1	S	Yes
64	c.1135C>T	p.(Arg379*)	Exon 8	SP	5	S	Yes
65	c.1136G>A	p.(Arg379Gln)	Exon 8	SP	2	Mo	Yes
66	c.1159_1161del	p.(Lys387del)	Exon 8	SP	1	S	No
67	c.1181T>A	p.(Met394Lys)	Exon 8	SP	1	S	Yes
68	c.1187G>A	p.(Cys396Tyr)	Exon 8	SP	1	S	Yes
69	c.1213G>A	p.(Asp405Asn)	Exon 8	SP	7	S	Yes
70	c.1220G>C	p.(Cys407Ser)	Exon 8	SP	1	S	No
71	c.1220G>A	p.(Cys407Tyr)	Exon 8	SP	1	S	Yes
72	c.1238G>A	p.(Gly413Glu)	Exon 8	SP	1	S	Yes
73	c.1243C>A	p.(His415Asn)	Exon 8	SP	2	S	Yes
74	c.1276A>G	p.(Thr426Ala)	Exon 8	SP	1	S	Yes
75	c.1277C>G	p.(Thr426Ser)	Exon 8	SP	1	M	Yes
76	c.1298A>C	p.(Glu433Ala)	Exon 8	SP	1	S	Yes
77	c.1298_1304del	p.(Glu433Valfs*2)	Exon 8	SP	1	Mo	No
78	c.1290C>A	p.(Ser430Arg)	Exon 8	SP	1	S	Yes
79	c.1331A>G	p.(Tyr444Cys)	Exon 8	SP	1	S	Yes
80	c.1369A>T	p.(Lys457*)	Exon 8	SP	3	S	Yes
81	c.1372dup	p.(Thr458Asnfs*18)	Exon 8	SP	1	S	Yes
82	c.865-?_2782+?del	p.(?)	Exon 7 to 8	SP	1	S	Yes
83	c.115-?_2782+?del	p.(?)	Exon 1 to 8	All	1	S	Yes

Abbreviations: HGVS, Human Genome Variation Society; S, Severe; Mo, Moderate; M, Mild; PRO, Pro-Peptide; AP, Activation Peptide; SP, Serine Protease; All, All domains; N/A, Not Applicable.

Identification of the Pathogenicity of Novel Variants

Our comprehensive molecular analysis characterized 20 novel variants in the F9 gene identified through systematic genetic screening and variant assessment. The variant spectrum comprised 6 missense mutations (30%), 3 nonsense mutations (15%), 6 small deletions/insertions (30%), 2 large deletions (10%), 2 splice site variants (10%), and 1 regulatory region variant (5%). The missense variants (p.Val30Phe, p.Cys64Trp, p.Cys157Phe, p.Gly230Val, p. Gly357Arg, p.Cys407Ser) demonstrated reduced FIX activity levels ranging from 0.1% to 0.5%, with in silico REVEL scores consistently predicting deleterious effects (0.854–0.976). All nonsense mutations (p.Glu193*p. Gln292*p.Lys362*) and frameshift variants (p.Glu116Glufs*15, p.Asn135*p.Leu318Phefs*8, p.Ser354Valfs*14, p. Glu433Valfs*2) resulted in severe phenotypes with FIX activity ≤ 0.5%. Two large genomic deletions were identified

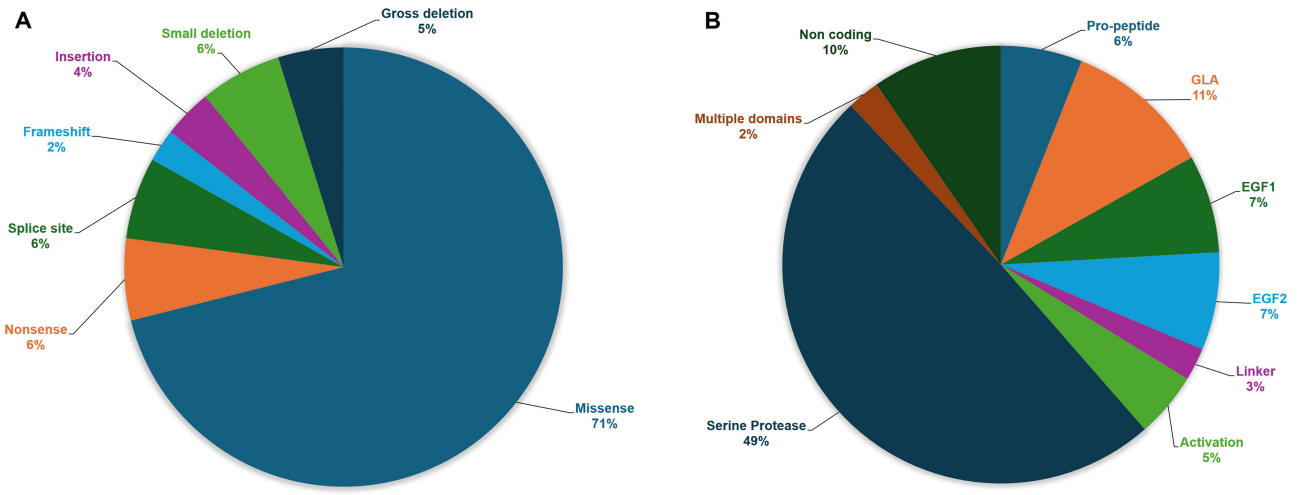


Figure 1 The frequency of 83 unique variants detected in 143 unrelated Vietnamese Hemophilia B patients classified by variant type (**A**) and by domain (**B**).

(c.547-?_749+?del and c.418-?_864+?del), both associated with severe disease manifestation. Splice site analysis revealed a significant predicted impact for c.502+2 T>C (SpliceAI DS=0.66), while the 5′ UTR variant c.-21C>T showed minimal predicted regulatory effect (SpliceAI DS=0.01). ACMG/AMP variant classification criteria were systematically applied to each variant, incorporating clinical phenotype data, population frequency, computational predictions, and functional domain impacts. This assessment resulted in the classification of 12 variants (60%) as Pathogenic, supported by PVS1 criterion and additional evidence (PM2, PP4); 6 variants (30%) as Likely Pathogenic, based on combinations of moderate (PM1, PM2, PM5) and supporting criteria (PP3, PP4); and 2 variants (10%) as Variants of Uncertain Significance (VUS), reflecting limited evidence for definitive classification. Factor IX activity measurements demonstrated correlation with variant classification, with Pathogenic variants typically associated with severe deficiency (<1%), while Likely Pathogenic variants generally corresponded to moderate deficiency (1–5%). The clinical classification of novel variants is described in Table 3.

Table 3 Pathogenicity Prediction of Novel Variants in F9 Gene

No	HGVS cDNA	HGVS Protein	FIX Level (%)	In Silico Prediction	ACMG Criteria	Predicted Clinical Significance
1	c.88G>T	p.(Val30Phe)	0.4	0.854	PM2 (1 pts), PP4 (2 pts), PP3 (1 pts), PM5 (2pts)	LP
2	c.192T>G	p.(Cys64Trp)	< 1%	0.932	PM2 (1 pts), PP4 (2 pts), PP3 (1 pts), PM5 (2pts)	LP
3	c.470G>T	p.(Cys157Phe)	0.2	0.972	PM1 (2 pts), PM2 (2 pts), PP3 (1 pts), PP4 (2 pts)	LP
4	c.577G>T	p.(Glu193*)	0.5	N/A	PVS1 (8 pts), PM2 (2pts), PP4 (1pts)	P
5	c.689G>T	p.(Gly230Val)	0.1	0.976	PM2 (2pts), PP3 (1pts), PP4 (1pts), PM5 (2 pts)	LP
6	c.874C>T	p.(Gln292*)	0.4	N/A	PVS1 (8 pts), PM2 (2pts), PP4 (1pts)	P
7	c.1069G>C	p.(Gly357Arg)	0.1	0.953	PM2 (1 pts), PP4 (2 pts), PP3 (1 pts), PS1 (4 pts), PM5 (2pts)	LP
8	c.1084A>T	p.(Lys362*)	0.1	N/A	PVS1 (8 pts), PM2 (2pts), PP4 (1pts)	P

(*Continued*)

Table 3 (Continued).

No	HGVS cDNA	HGVS Protein	FIX Level (%)	In Silico Prediction	ACMG Criteria	Predicted Clinical Significance
9	c.1220G>C	p.(Cys407Ser)	0.5	0.973	PM2 (2 pts), PP4 (2 pts), PP3 (1 pts), PS1 (2pts), PM5 (2pts)	LP
10	c.348_349insGA	p.(Glu116Glufs*15)	0.3	N/A	PVS1 (8 pts), PM2 (2pts), PP4 (1pts)	P
11	c.407_408insGTAACAT	p.(Asn135*)	0.2	N/A	PVS1 (8 pts), PM2 (2pts), PP4 (1pts)	P
2	c.929del	p.(Asn210Ilefs*15)	0.1	N/A	PVS1 (8 pts), PM2 (2pts), PP4 (1pts)	P
13	c.952del	p.(Leu318Phefs*8)	0.1	N/A	PVS1 (8 pts), PM2 (2pts), PP4 (1pts)	P
14	c.1060del	p.(Ser354Valfs*14)	0.3	N/A	PVS1 (8 pts), PM2 (2pts), PP4 (1pts)	P
15	c.1159_1161del	p.(Lys387del)	0.2	0.85	PM2 (2 pts), PM4 (2pts), PP4 (2pts)	VUS
16	c.1298_1304del	p.(Glu433Valfs*2)	4.1	N/A	PVS1 (8 pts), PM2 (2pts), PP4 (1pts)	P
17	c.547-?_749+?del	p.(?)	0.4	N/A	PVS1 (8 pts), PM2 (2pts), PP4 (1pts)	P
18	c.418-?_864+?del	p.(?)	< 1%	N/A	PVS1 (8 pts), PM2 (2pts), PP4 (1pts)	P
19	c.-21C>T	N/A	4.5	DS = 0.01	PM2 (2pts), PP4 (1pts)	VUS
20	c.502+2 T>C	N/A	0.1	DS = 0.66	PVS1 (8 pts), PM2 (2pts), PP3 (1 pts), PP4 (1pts)	P

Abbreviations: HGVS, Human Genome Variation Society; FIX, Factor IX; P, Pathogenic; LP, Likely Pathogenic; VUS, Variant of Uncertain Significance; N/A, Not Applicable; pts, points.

Discussion

Our research is the first large-scale study on the molecular genetics of HB in Vietnam, offering critical insights into the genetic and clinical characteristics of HB patients. The predominance of male patients in the cohort is consistent with the X-linked inheritance pattern of HB, which results in the disease primarily affecting males, while females typically act as carriers.[12] Symptomatic females are quite rare.[13] The identification of a female patient with HB represents a clinically significant and genetically complex scenario. Analysis of the inheritance pattern reveals two potential molecular mechanisms: either the inheritance of pathogenic F9 variants from both parents (compound heterozygosity), or the inheritance of a single paternal variant accompanied by non-random X-chromosome inactivation (skewed X–inactivation). Molecular characterization revealed the patient carries the c.881G>A variant in homozygous state, with comprehensive family analysis confirming inheritance from a hemizygous father (c.881G>A) and a heterozygous carrier mother (c.881G>A). This unusual inheritance pattern, while statistically uncommon, can emerge in populations with elevated rates of consanguinity or in geographically isolated communities where the frequency of hemophilia B carriers and affected individuals may be higher. This case underscores the importance of comprehensive genetic analysis and pedigree evaluation in understanding atypical presentations of X-linked disorders, particularly in populations with unique demographic characteristics.

This comprehensive molecular analysis achieved complete variant detection in the *F9* gene across all 143 Vietnamese patients with HB, identifying 83 unique variants. This detection rate aligns with findings from recent large-scale studies, including the work by Goodeve et al documenting >97% mutation detection rates using current molecular diagnostic approaches.[14] The variant spectrum revealed a predominance of SNVs (85.5%), most of which were missense variants. This distribution parallels findings from the international FIX Variant Database (F9db), which documented missense variants as the most frequent mutation type among >3,000 pathogenic variants.[15] Among these, 71.08% were missense variants affecting critical functional domains, consistent with Rallapalli et al's (2013) analysis of structure-function relationships in FIX.[6] Small deletions/insertions occurred at lower frequencies (8.43%), comparable to rates reported in

the studies of Huang et al (2020) and Wang et al (2016).[16,17] Large deletions, affecting one or multiple exons, account for 6.02% of the variants, consistent with Li et al's (2014) comprehensive analysis of HB populations in the US.[18]

Analysis of affected domains reveals that the protease domain is the most frequently impacted (49.4%), showing its central role in maintaining FIX activity and its association with severe phenotypes. This finding is consistent with those of previous reports, which highlight the protease domain as a hotspot for pathogenic variants.[19–21] The considerable involvement of the pro-peptide (6.02%) and GLA (10.84%) domains supports their important roles in protein stabilization and activation, corroborating findings by Shen et al.[22] Variants in the EGF1 (7.23%) and EGF2 (7.23%) domains emphasize their role in structural integrity and receptor interactions, further emphasizing the functional importance of these regions. Rarely affected regions, such as the linker domain (2.41%), promoter (2.41%), and multiple domains (2.41%), indicate less-characterized roles in disease but indicate potential contributions to specific phenotypic outcomes. The lower prevalence of variants in these regions aligns with their less critical roles compared to catalytic and structural domains.[13]

In an interesting case, the detection of the c.-21C>T variant in the *F9* gene aligns with HB Leyden, a rare subtype of HB characterized by age-related changes in FIX levels due to androgen-regulated transcriptional activation.[23,24] Located in the promoter region, this variant disrupts the binding transcription factors necessary for FIX expression during early life.[25] The gradual increase in FIX levels observed in this patient, from 4% in the prepubertal phase to 8.9% at age 14 and 23.6% at age 18, is consistent with the typical improvement in FIX levels seen in HB Leyden post-puberty.[24,26] This biphasic pattern underscores a need for tailored management, with more intensive prophylaxis in early childhood and the potential for reduced therapy or cessation in adulthood as FIX levels naturally increase. The case also highlights the importance of understanding the molecular mechanisms of *F9* promoter variants, which may inform therapeutic strategies such as hormonal modulation or gene editing to enhance FIX production.

The identification of three *F9* variants exhibiting mixed severity patterns (c.128G>A, c.127C>T, and c.880C>T) provides valuable insights into the complex genotype-phenotype relationships in HB. These findings align with Chavali et al's observation of "phenotypic plasticity" in HB, where certain mutations show variable clinical manifestations.[27] The variants c.128G>A and c.127C>T, both affecting arginine 43 in the pro-peptide domain, demonstrated severity ranging from moderate to severe. This phenotypic variability is particularly notable given the pro-peptide domain's crucial role in protein processing and secretion. Goodeve's comprehensive review emphasizes that variants affecting conserved regions can have varying functional impacts depending on the specific amino acid substitution.[14] The nonsense variant c.880C>T (p.Arg294*) showed an unexpected moderate to severe phenotype pattern. This finding may be explained by Pinotti et al's work on ribosome readthrough in HB, where certain nonsense mutations can allow minimal full-length protein production, potentially moderating the clinical severity.[28] The variable phenotype could also reflect the influence of genetic modifiers or environmental factors, as suggested by Santagostino et al's study on patients with severe hemophilia and mild bleeding tendencies.[29] These mixed severity patterns underscore the limitations of traditional factor activity-based classification systems. As demonstrated by Mancuso et al's work on clinically severe hemophilia, factor levels alone may not fully predict the bleeding phenotype.[30] The identification and characterization of such variants with variable expressivity contribute to our understanding of the molecular pathogenesis of HB and highlight the importance of considering genetic and clinical parameters in patient assessment and management. Future functional studies of these variants could provide valuable insights into the mechanisms underlying phenotypic variability in HB.

This comprehensive analysis of 20 novel variants in the *F9* gene provides crucial insights into the molecular pathogenesis of Hemophilia B and advances our understanding of variant interpretation in clinical settings. Using the standardized ACMG/AMP guidelines, our study revealed a diverse spectrum of mutations including nonsense, missense, insertions/deletions, and splice site alterations, with 12 variants (60%) classified as Pathogenic, 6 (30%) as Likely Pathogenic, and 2 (10%) as Variants of Uncertain Significance.[10] Nine loss-of-function variants resulting in premature termination codons were uniformly classified as Pathogenic based on very strong evidence (PVS1), supported by extensive functional studies demonstrating their role in severe disease through haploinsufficiency.[31] The six missense variants, predominantly classified as Likely Pathogenic, showed strong correlation with critical functional domains and demonstrated consistent in silico predictions with REVEL scores ranging from 0.854 to 0.976, aligning with recent validation studies.[32] Of particular interest was the identification of a 5′ UTR variant (c.-21C>T) classified as VUS,

highlighting the ongoing challenges in interpreting regulatory region variants, especially given recent studies demonstrating complex mechanisms of F9 transcriptional regulation.[33] Notably, our analysis revealed genotype-phenotype correlations, with Pathogenic variants generally associated with severe deficiency (<1% FIX activity) and Likely Pathogenic variants corresponding to moderate deficiency (1–5%), though individual variation exists as previously documented.[14] The robust classification system employed in this study demonstrates the utility of standardized variant interpretation while simultaneously identifying areas requiring additional evidence, particularly in the functional validation of VUS and regulatory variants. Our findings support the continued refinement of F9-specific interpretation guidelines and emphasize the critical importance of integrating multiple lines of evidence for accurate pathogenicity assessment, ultimately contributing to improved diagnostic accuracy and clinical management of Hemophilia B patients.

Based on the analyzed variant data, a comprehensive mapping of *F9* gene variants was constructed for HB patients included in this study. The variant map illustrates all 83 identified variants, their locations, frequencies across exons and introns of the *F9* gene, and their distribution within the functional domains of the FIX molecule (Figure 2). This map provides a global characterization of the genetic alterations in the *F9* gene among Vietnamese patients with HB.

The wide variety of genetic variants identified in this study, including missense, nonsense, frameshift, splice-site mutations, large deletions, and regulatory changes, underscores the complexity of the molecular background of the disease. These variants contribute to a spectrum of disease severities, ranging from mild to severe, depending on their impact on FIX protein synthesis, function, or regulation. While loss-of-function mutations, such as nonsense and frameshift variants, are often associated with severe phenotypes due to significantly reduced FIX activity, missense mutations and regulatory variants may result in variable phenotypes based on residual protein function or compensatory mechanisms, as seen in cases like HB Leyden.[23,26] Despite these established correlations, several variants remain classified as VUS, highlighting the gaps in our understanding of their functional impact. Future studies should investigate the correlation between genotype and phenotype to refine the pathogenicity of these variants, improve the accuracy of genetic diagnoses, and personalize treatment strategies. Such studies should integrate functional assays, structural modeling, and large-scale genotype-phenotype datasets to deepen our understanding of the clinical diversity observed in HB. Finally, carrier testing and genetic studies in female relatives are crucial to identify potential carriers and predict disease inheritance patterns, especially in cases where severe variants are present in a male member of the family.

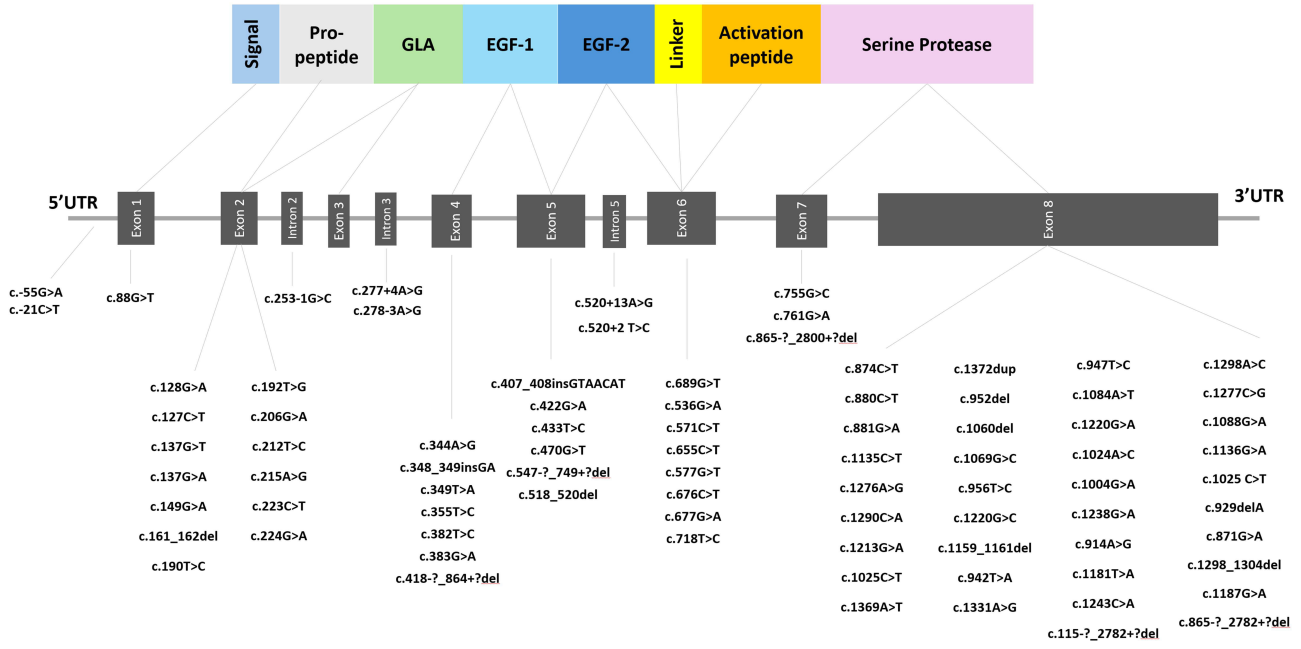


Figure 2 Distribution map of 83 unique variants of 143 unrelated Vietnamese patients with Hemophilia B. Variants were distributed across the coding, non-coding, and untranslated regions. Most variants were located in the coding regions spanning exons 1 to 8, with exon 8 exhibiting the highest number of variants (n = 35), followed by exon 2 (n = 13). Six intronic variants were identified, and variants in the 5′ UTR were observed in three patients.

https://doi.org/10.2147/JBM.S514338

Conclusion

This study reinforces the established mutation spectrum in HB while providing a deeper understanding of the structural and functional impacts of variants in the *F9* gene. These findings highlight the importance of implementing comprehensive genomic analyses in Vietnam, including advanced sequencing technologies, to capture the diverse mutational landscape comprehensively. Future studies should further investigate the functional consequences of rare variants and explore novel therapeutic interventions targeting specific domains and variant types.

Abbreviations

HB, Hemophilia B; FIX, Factor IX; DNA, Deoxyribonucleic Acid; CDC, Centers for Disease Control and Prevention; EAHAD, European Association for Haemophilia and Allied Disorders; NCBI, National Center for Biotechnology Information; ACMG, American College of Medical Genetics and Genomics; AMP, Association for Molecular Pathology; HGVS, Human Genome Variation Society; MLPA, Multiplex Ligation-Dependent Probe Amplification; SNV, Single Nucleotide Variants; VUS, Variant of Uncertain Significance.

Data Sharing Statement

The data that support the findings of this study are available from the corresponding author, [THN], upon reasonable request.

Ethical Approval and Informed Consent

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the National Institute of Hematology and Blood Transfusion, Hanoi, Vietnam (protocol code 2525/QD-HHTM, date of approval: 20/12/2018). Informed consent was obtained from all subjects involved in the study.

Consent for Publication

Written informed consent for publication of their details was obtained from the study participants and parent or legal guardian of patients under 18 years of age.

Acknowledgments

The authors would like to express sincerest thanks to the Ministry of Science and Technology of Vietnam, for their financial support to this study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was funded by the Ministry of Science and Technology of Vietnam (code: NDT/AU/21/26).

Disclosure

Dr Khanh Bach reports grants from Vietnam Ministry of Science and Technology, during the conduct of the study. Dr Chinh Duong reports grants from Vietnam Ministry of Science and Technology, during the conduct of the study. Dr Huong Vu reports grants from Vietnam Ministry of Science, during the conduct of the study. Dr Binh Nguyen reports grants from Vietnam Ministry of Science, during the conduct of the study. Mrs Trang Nguyen reports grants from Vietnam Ministry of Science, during the conduct of the study. Dr Mai Nguyen reports grants from Vietnam Ministry of

Science and Technology, during the conduct of the study. Prof. Dr. Thanh Nguyen reports grants from Vietnam Ministry of Science and Technology, during the conduct of the study. The authors report no conflicts of interest in this work.

References

1. Srivastava A, Santagostino E, Dougall A, et al. WFH Guidelines for the Management of Hemophilia. *Haemophilia*. 2020;26(Suppl 6):1–158. doi:10.1111/hae.14046
2. Stonebraker JS, Bolton-Maggs PHB, Soucie JM, Walker I, Brooker M. A study of variations in the reported haemophilia A prevalence around the world. *Haemophilia*. 2010;16(1):20–32. doi:10.1111/j.1365-2516.2009.02127.x
3. Angchaisuksiri P, Amurao-Abiera M, Chou SC, et al. Haemophilia care in Asia: learning from clinical practice in some Asian countries. *Haemophilia*. 2024;30(3):609–616. doi:10.1111/hae.14998
4. Thi Mai N. Hemophilia Care in Vietnam. Ministy of Health. National Institute of Hematology and Blood Transfusion. Available from : https://ipfa.nl/wp-content/uploads/2018/10/2_2_-NGUYEN-MAI-IPFA_hanoi-3-ngay-05.3.2019-da-sua-2.pdf.
5. Giannelli F, Green PM, Sommer SS, et al. Haemophilia B (sixth edition): a database of point mutations and short additions and deletions. *Nucleic Acids Res*. 1996;24(1):103–118. doi:10.1093/nar/24.1.103
6. Rallapalli PM, Kemball-Cook G, Tuddenham EG, Gomez K, Perkins SJ. An interactive mutation database for human coagulation factor IX provides novel insights into the phenotypes and genetics of hemophilia B. *J Thromb Haemost*. 2013;11(7):1329–1340. doi:10.1111/jth.12276
7. Pruthi RK. Hemophilia: a Practical Approach to Genetic Testing. *Mayo Clin Proc*. 2005;80(11):1485–1499. doi:10.4065/80.11.1485
8. Mitchell M, Keeney S, Goodeve A. The molecular analysis of haemophilia B: a guideline from the UK haemophilia centre doctors' organization haemophilia genetics laboratory network. *Haemophilia*. 2005;11(4):398–404. doi:10.1111/j.1365-2516.2005.01112.x
9. Untergasser A, Cutcutache I, Koressaar T, et al. Primer3–new capabilities and interfaces. *Nucleic Acids Res*. 2012;40(15):e115. doi:10.1093/nar/gks596
10. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405–424. doi:10.1038/gim.2015.30
11. Tavtigian SV, Harrison SM, Boucher KM, Biesecker LG. Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines. *Hum Mutat*. 2020;41(10):1734–1737. doi:10.1002/humu.24088
12. White GC, Rosendaal F, Aledort LM, et al. Definitions in hemophilia. Recommendation of the scientific subcommittee on factor VIII and factor IX of the scientific and standardization committee of the International Society on Thrombosis and Haemostasis. *Thromb Haemost*. 2001;85(3):560. doi:10.1055/s-0037-1615621
13. Plug I, Mauser-Bunschoten EP, Bröcker-Vriends AHJT, et al. Bleeding in carriers of hemophilia. *Blood*. 2006;108(1):52–56. doi:10.1182/blood-2005-09-3879
14. Goodeve AC. Hemophilia B: molecular pathogenesis and mutation analysis. *J Thromb Haemost*. 2015;13(7):1184–1195. doi:10.1111/jth.12958
15. McVey JH, Rallapalli PM, Kemball-Cook G, et al. The European Association for Haemophilia and Allied Disorders (EAHAD) Coagulation Factor Variant Databases: important resources for haemostasis clinicians and researchers. *Haemophilia*. 2020;26(2):306–313. doi:10.1111/hae.13947
16. Wang QY, Hu B, Liu H, et al. A genetic analysis of 23 Chinese patients with hemophilia B. *Sci Rep*. 2016;6(1):25024. doi:10.1038/srep25024
17. Huang L, Li L, Lin S, et al. Molecular analysis of 76 Chinese hemophilia B pedigrees and the identification of 10 novel mutations. *mol Genet Genomic Med*. 2020;8(11):e1482. doi:10.1002/mgg3.1482
18. Li T, Miller CH, Driggers J, Payne AB, Ellingsen D, Hooper WC. Mutation analysis of a cohort of US patients with hemophilia B. *Am J Hematol*. 2014;89(4):375–379. doi:10.1002/ajh.23645
19. Liu G, Sun J, Li Z, Chen Z, Wu W, Wu R. F9 mutations causing deletions beyond the serine protease domain confer higher risk for inhibitor development in hemophilia B. *Blood*. 2023;141(6):677–680. doi:10.1182/blood.2022017871
20. Kulkarni S, Hegde R, Hegde S, et al. Mutation analysis and characterisation of F9 gene in haemophilia-B population of India. *Blood Res*. 2021;56(4):252–258. doi:10.5045/br.2021.2021016
21. Ghosh K, Quadros L, Shetty S. Spectrum of factor IX gene mutations causing haemophilia B from India. *Blood Coagulation & Fibrinolysis: an International Journal in Haemostasis and Thrombosis*. 2009;20(5):333–336. doi:10.1097/MBC.0b013e32832b27d1
22. Shen G, Gao M, Cao Q, Li W. The Molecular Basis of FIX Deficiency in Hemophilia B. *Int J mol Sci*. 2022;23(5):2762. doi:10.3390/ijms23052762
23. Crossley M, Winship PR, Austen DE, Rizza CR, Brownlee GG. A less severe form of Haemophilia B Leyden. *Nucleic Acids Res*. 1990;18(15):4633. doi:10.1093/nar/18.15.4633
24. Kurachi S, Huo JS, Ameri A, Zhang K, Yoshizawa AC, Kurachi K. An age-related homeostasis mechanism is essential for spontaneous amelioration of hemophilia B Leyden. *Proc Natl Acad Sci U S A*. 2009;106(19):7921–7926. doi:10.1073/pnas.0902191106
25. Funnell APW, Wilson MD, Ballester B, et al. A CpG mutational hotspot in a ONECUT binding site accounts for the prevalent variant of hemophilia B Leyden. *Am J Hum Genet*. 2013;92(3):460–467. doi:10.1016/j.ajhg.2013.02.003
26. Beskorovainaya TS, Zabnenkova VV, Zinchenko RA, Shchagina OA, Polyakov AV. Hemophilia B Leyden: literature and Our Data. *Russ J Genet*. 2021;57(10):1131–1139. doi:10.1134/S1022795421100033
27. Chavali S, Ghosh S, Bharadwaj D. Hemophilia B is a quasi-quantitative condition with certain mutations showing phenotypic plasticity. *Genomics*. 2009;94(6):433–437. doi:10.1016/j.ygeno.2009.08.005
28. Pinotti M, Caruso P, Canella A, et al. Ribosome readthrough accounts for secreted full-length factor IX in hemophilia B patients with nonsense mutations. *Human Mutation*. 2012;33(9):1373–1376. doi:10.1002/humu.22120
29. Santagostino E, Mancuso ME, Tripodi A, et al. Severe hemophilia with mild bleeding phenotype: molecular characterization and global coagulation profile. *J Thromb Haemost*. 2010;8(4):737–743. doi:10.1111/j.1538-7836.2010.03767.x
30. Mancuso ME, Bidlingmaier C, Mahlangu JN, Carcao M, Tosetto A. subcommittee on factor viii, factor ix and rare coagulation disorders. The predictive value of factor VIII/factor IX levels to define the severity of hemophilia: communication from the SSC of ISTH. *J Thromb Haemost*. 2018;16(10):2106–2110. doi:10.1111/jth.14257

31. Peyvandi F, Kunicki T, Lillicrap D. Genetic sequence analysis of inherited bleeding diseases. *Blood*. 2013;122(20):3423–3431. doi:10.1182/blood-2013-05-505511
32. Harrison SM, Biesecker LG, Rehm HL. Overview of Specifications to the ACMG/AMP Variant Interpretation Guidelines. *Curr Protoc Hum Genet*. 2019;103(1):e93. doi:10.1002/cphg.93
33. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell*. 2019;176(3):535–548.e24. doi:10.1016/j.cell.2018.12.015